

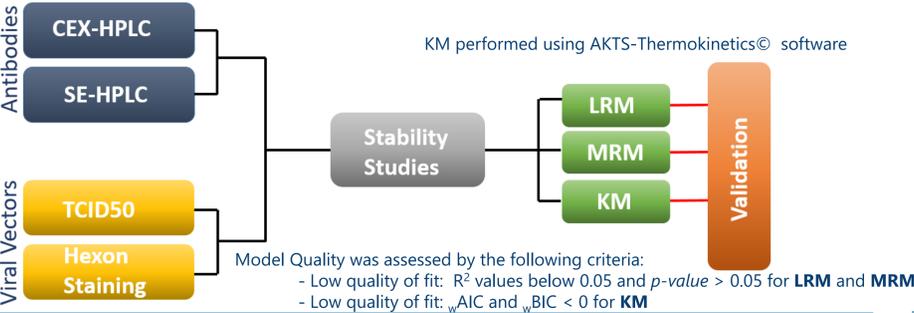


### Introduction

- Estimating shelf-life of drug products is crucial to ensure efficacy, purity and potency
  - Shelf life is estimated from long term stability and accelerated aging studies
  - The classical approach approved by ICH guidelines is to perform **linear regression** on stability attributes, on the condition that the degradation pattern is constant
  - Previous research show that **alternative methods** to linear regression could better tackle the degradation of **antibodies [1, 2] and viral vectors [3, 4]** which exhibit complex degradation patterns composed of multiple steps and autocatalytic behaviour, leading to **superior stability predictions for selection candidates**
- Here we compared the accuracy of **three predictive methods** on the stability of 2 antibodies and 2 viral vectors **after storage at 5°C and 25°C for 6 months** : linear (LRM), multiple regression (MRM) and kinetic modelling (KM).

### Research Questions and Methods

- How accurate are predictions using data up to 3m compared to measured data at 6m?
- How do predicted results compare between the three modelling methods?
- What is the minimum of time points necessary for accurate predictions?



3. What is the minimum of time points necessary for accurate predictions?

**Plus:** Is there a modelling method working best for mAb vs. viral vectors?

Temperature	Time points Used for Data Generation						
	0 weeks	1 week	2 week	1 month	2 months	3 months	6 months
5°C	X			X	X	X	X
25°C	X			X	X	X	X
40°C	X	X	X	X			

- Dataset 1: Using all data up to 1 month for three temperatures
- Dataset 2a: Using all data up to 2 months for three temperatures
- Dataset 2b: Using all data up to 2 months for 5°C, 25°C, and only t = 0, and t = 1 month data for 40°C
- Dataset 3a: Using all data up to 3 months for three temperatures
- Dataset 3b: Using all data up to 3 months for 5°C, 25°C, and only t = 0, and t = 1 month data for 40°C

### Results

#### CEX-HPLC storage at 5°C

Modelling Method* (Dataset 3a)	Main Peak Area [%]	
	Error [Measured - Predicted] mAb A	mAb B
LRM	0.56	-0.52
MRM	0.56	-0.52
KM	0.52	-0.71

#### SE-HPLC storage at 5°C

Modelling Method* (Dataset 3a)	Monomer Peak Area [%]	
	Error [Measured - Predicted] mAb A	mAb B
LRM	0.01	0.16
MRM	0.01	0.16
KM	-0.11	-0.16

#### Hexon Staining storage at 5°C and 25°C

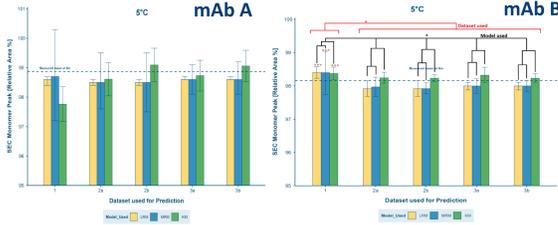
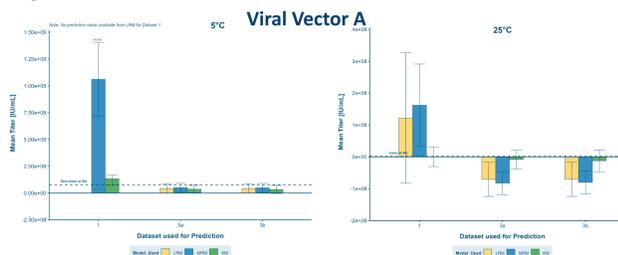
Modelling Method* (Dataset 3a)	Mean Titer [IU/ml]	
	Error [Measured - Predicted] (Error %) 5°C	25°C
LRM	3.38E+07 (44.14%)	7.22E+07 (4204.56%)
MRM	2.70E+07 (35.21%)	8.41E+07 (4898.23%)
KM	4.35E+07 (56.81%)	1.41E+07 (824.19%)

All models for both antibodies show less than 1% prediction error from measured stability at 6 months.

All models for both antibodies less than 1% prediction error from measured stability at 6 months.

KM revealed more reliable predictions of Mean Titer at 25°C.

\* Most LRM models for stability at 5°C and 25°C showed low quality of fit and were removed from analyses (no ANOVA possible). At 25°C, LM and MRM expresses very high variation (see CI 95%).



KM revealed more reliable predictions of Mean Titer at 25°C.

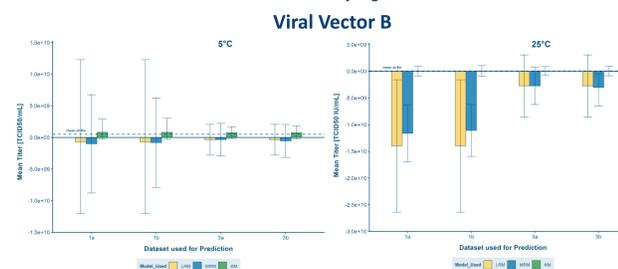
For datasets with 3 months data, all methods were similarly accurate in predicting measured stability at 6 months. (One-Sample t-test against mean Titer,  $p < 0.038$  for Dataset with 1 month data and MRM)

#### TCID50 storage at 5°C and 25°C

Modelling Method (Dataset 3a) *	Mean Titer [TCID50/ml]	
	Error [Measured - Predicted] (Error %) 5°C	25°C
LRM	8.60E+08 (170.61%)	2.80E+09 (3783.41%)
MRM	8.60E+08 (170.61%)	2.80E+09 (3783.41%)
KM	-2.52E+08 (49.91%)	1.19E+07 (16.03%)

KM revealed more reliable predictions of Mean Titer at 25°C.

\* Most LRM models for stability at 5°C and 25°C showed low quality of fit and were removed from analyses (no ANOVA possible). Predictions were on negative scale for LRM and MRM. At 25°C, LM and MRM revealed very high variation.



KM was the only method accurate enough to be compared to measured Titer, as seen by CI bars. (One-Sample t-test against mean Titer,  $p > 0.05$  for all methods and datasets)

### Discussion

Current ICH guidelines support the use of Linear Regression as golden standard to predict shelf life stability of drug products. In agreement with previous findings [1,2] kinetic modelling showed to be highly accurate in predicting degradation pathways of standard antibodies as measured by HPLC methods. Nevertheless, **linear and multiple regression models showed similar results in accuracy, supporting ICH suggestions.**

However, **shortcomings of linear regression** become evident when assays with high variability and complex degradation pathways are under analysis. **Kinetic modelling showed to be the only method able to handle higher temperature stresses [3, 4]** and provide meaningful insights in viral vectors degradation.

### Summary

Stability prediction accuracy is **highly dependent on measurement** assays and drug substances. **Antibodies** measured with CEX and SEC showed **consistent results**, while **viral vectors** as measured by Hexon Staining and TCID50 presented a huge **modelling challenge**.

#### Antibody stability

- All predictions at 5°C and 25°C for CEX and SEC with training data up to 3 months showed **very high accuracy** compared to measured stability at 6 months
- All modelling methods were interchangeable in providing accurate results, but **Kinetic Modelling showed to be more accurate with increasing storage temperature**
- There is **no need to collect more than 2 months data** to predict 6 months stability for standard mAbs

#### Viral stability

- Predictions at 5°C and 25°C for TCID50 and Hexon Staining show degradation irregularities between training data up to 3 months and measured stability at 6 months which increased inaccuracy
- At 5°C storage, all modelling methods revealed inaccurate predictions. With increasing temperature, **at 25°C Kinetic Modelling showed consistently more accurate predictions** compared to measured stability at 6 months. Oppositely, linear regression consistently showed the lowest prediction outcome
- A **minimum of 3 months data** is strongly suggested to infer predictions for 6 months stability

### References

- Kuzman, D., Bunc, M., Ravnik, M., Reiter, F., Žagar, L., & Bončina, M. (2021). Long-term stability predictions of therapeutic monoclonal antibodies in solution using Arrhenius-based kinetics. *Scientific Reports*, 11
- Evers, A., Clénet, D., & Pfeiffer-Marek, S. (2022). Long-Term Stability Prediction for Developability Assessment of Biopharmaceuticals Using Advanced Kinetic Modeling. *Pharmaceutics*, 14(2), 375
- Clénet, D. (2018). Accurate prediction of vaccine stability under real storage conditions and during temperature excursions. *European Journal of Pharmaceutics and Biopharmaceutics*, 125, 76–84.
- Campa, C., Ponce, T., Paludi, M., Weusten, J., Conway, L., Savery, J., Richards, C., & Clénet, D. (2021). Use of Stability Modeling to Support Accelerated Vaccine Development and Supply. *Vaccines*, 9(10), 1114. <https://doi.org/10.3390/vaccines9101114>

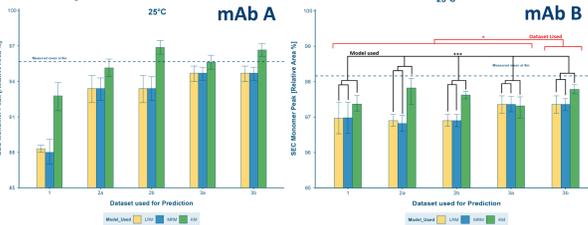
#### CEX-HPLC storage at 25°C

Modelling Method* (Dataset 3a)	Main Peak Area [%]	
	Error [Measured - Predicted] mAb A	mAb B
LRM	0.52	6.54
MRM	0.52	6.56
KM	0.55	2.91

#### SE-HPLC storage at 25°C

Modelling Method* (Dataset 3a)	Monomer Peak Area [%]	
	Error [Measured - Predicted] mAb A	mAb B
LRM	0.89	0.50
MRM	0.89	0.50
KM	0.59	0.54

All models for both antibodies less than 1% prediction error from measured stability at 6 months.



Prediction accuracy at 6 months increases with increasing time points in the dataset and particularly with KM compared to LM and MRM.

(Two-way ANOVA significant main effect of Dataset, Model Used and their Interaction ( $p < 0.05$ ) confirmed by Post hoc Tukey's HSD Test,  $p < 0.001$  for Dataset,  $p < 0.001$  for Model used and  $p < 0.011$  for their Interaction)

Datasets with 3 months data showed to be more accurate in predicting 6 months data. KM revealed to be significantly more accurate in comparison to LM and MRM.

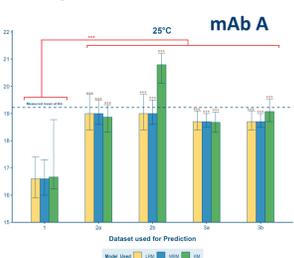
(Two-way ANOVA significant main effect of Dataset and Model Used ( $p < 0.05$ ) confirmed by Post hoc Tukey's HSD Test,  $p < 0.05$  for Dataset,  $p < 0.001$  for Model Used)

Any dataset with 2 months or more data revealed to be significantly more accurate in predicting 6 months stability independently on the modelling methods used. (Two-way ANOVA significant main effect of Dataset ( $p < 0.05$ ), confirmed by Post hoc Tukey's HSD Test,  $p < 0.001$ )

#### CEX-HPLC storage at 25°C

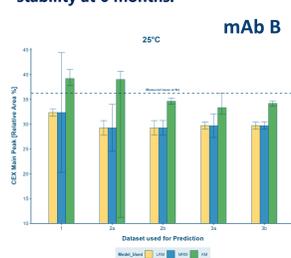
Modelling Method* (Dataset 3a)	Main Peak Area [%]	
	Error [Measured - Predicted] mAb A	mAb B
LRM	0.52	6.54
MRM	0.52	6.56
KM	0.55	2.91

mAb A: All models A showed less than 1% prediction error from measured stability at 6 months.



Any dataset with 2 months or more data revealed to be significantly more accurate in predicting 6 months stability independently on the modelling methods used. (Two-way ANOVA significant main effect of Dataset ( $p < 0.05$ ), confirmed by Post hoc Tukey's HSD Test,  $p < 0.001$ )

mAb B: KM showed the lowest prediction error from measured stability at 6 months.



No effect detected of dataset, modelling method or their interactions. (Two-way ANOVA) \*All models showed high quality of fit.